DOCUMENT RESUME

ED 178 598 TM 009 805

AUTHOR Mazzuca, Steven A.; Feldhusen, John F.
TITLE Causal Analysis of the Effects of Situational Variables on Student Evaluation of Instruction.
PUB DATE 9 Apr 79
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Biographical Inventories; Class Size; College Majors; *Course Evaluation; Critical Path Method; Grades (Scholastic); Higher Education; Hypothesis Testing; Participant Satisfaction; *Predictor Variables; Program Attitudes; *Student Characteristics; *Student Evaluation of Teacher Performance; Student Interests; *Student Motivation
IDENTIFIERS Causal Analysis

ABSTRACT

A theory, which proposed a hierarchy of determinants of student ratings of instruction, was tested. Situational variables such as class size, elective enrollment, and grade expectancy were hypothesized to be only indicators of more potent determiners. These motivational constructs were diversity of enrollment motives and subject matter interests, and satisfaction of important subject-related interests. The theory was tested using path analysis with a sample of 1,939 college students in 52 classes, taught by 43 instructors. Results supported the concept of a hierarchy of determinants. (Conclusions regarding the importance of study variables, and implications for the administrative and diagnostic use of ratings were addressed). (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED178598

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Causal Analysis of the Effects of Situational Variables on Student Evaluation of Instruction

Steven A. Mazzuca

Indiana University School of Medicine

John F. Feldhusen

Purdue University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Steven Mazzuca

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the Annual Meeting of the American Educational Research Association, April 9, 1979.

TM009 805

Causal Analysis of the Effects of Situational Variables on Student Evaluation of Instruction

by

Steven A. Mazzuca
Indiana University
School of Medicine

John F. Feldhusen
Purdue University

ABSTRACT

The purpose of this study was to test a theory which proposed a hierarchy of determinants of student ratings. Situational variables such as class size, elective enrollment, and grade expectancy were hypothesized to be only indicators of more potent determiners. These motivational constructs were diversity of enrollment motives and subject matter interests and satisfaction of important subject-related interests. The theory was tested using path analysis with a sample of 52 college classes (N = 1939 students). Results support the concept of a hierarchy of determinants. Conclusions regarding the importance of study variables and implications for the administrative and diagnostic use of ratings are addressed.

Causal Analysis of the Effects of Situational Variables on Student Evaluation of Instruction

by

Steven A. Mazzuca
Indiana University
School of Medicine

John F. Feldhusen
Purdue University

Numerous studies have investigated correlates of student ratings among various situational and student demographic variables. Among these variables are class size, academic major, academic level, elective versus required enrollemnt, grade expectations and the like. Reviews by Costin, Greenough, and Menges (1971) and Kulik and McKeachie (1975) suggested that correlational studies of the relationships among these situational variables and student ratings offer mixed findings.

Studies with statistically significant results generally associate high ratings with small class size (Frey, 1977; Heilman and Armentrout, 1936; Lovell and Haner, 1955; McDaniel and Feldhusen, 1970), high academic levels (Clark and Keller, 1954: Downie, 1952), elective enrollment (Cohen and Humphreys, 1960; Feldhusen and Mazzuca, in press) and high achievement or grade expectation (Rayder, 1968; Treffinger and Feldhusen, 1970). Nevertheless, Costin, et al. (1971) noted that, in almost all cases, correlations between ratings and these variables are at best moderate - generally in the range of .20 to .40.

Two sorts of conclusions might reasonably be drawn from these findings. The first, most emphatically asserted by Kulik and McKeachie (1975), is that student variables such as sex, age, grade, and major are of trivial importance as determinants of variation in student ratings of instruction. This is no doubt true, given the current, literal interpretation of situational variables. However, an alternate conclusion is that, given the

moderate correlations between situational variables and student ratings (four to sixteen percent common variance), these variables are really indicators of more potent determiners of rating variation.

Situational variables examined in the vast majority of studies in this area have been restricted to those which are obtainable directly from a rating form. Perhaps it is easy to see why these variables have been so popular. The question must be asked, however, "Just what do variables like class size, major, expected grade, and so forth actually represent?" Take, for example, class size. On the surface, this variable is nothing more than a "head count." But it certainly connotes more than just the number of seats occupied. In fact, it connotes too much. Depending on one's inclination, class size can indicate the amount of time available to each student for individual consultation, or it can dictate the most propitious methods of instruction and evaluation. There is no simple, readily apparent rationale for a relationship between the number of students who complete a rating form and the tendency of the general evaluation.

Similar concerns arise for the other variables. Of what theoretical significance are academic level or major of students, grade expectancy, and elective versus required enrollment? To date, research literature offers little in the way of answers. Investigators' major concern has been to tell users, "Yes, student evaluation of instruction is relatively free from these biases," or, "No, it is not."

The purpose of this research was to test a conceptual interpretation for the relationships among situational variables and student evaluation of

instruction. A diagrammatic representative of the theory appears in Figure 1. A brief description of the proposed interpreted framework is presented next.

Aside from students' ability to rate instructor and course characteristics objectively, student evaluations are in part determined by certain motivational constructs: (1) the reasons for their enrollment in particular courses and even sections, (2) specific interests in course material, and (3) the extent to which they perceive that these interests have been met, challenged, or otherwise satisfied. Large classes tend to be characterized by a wide variety of academic majors as well as diversity in academic levels (i.e., freshman, sophomores, juniors, etc.). Such heterogeneity is likely to result in a similar, underlying diversity in enrollment motives and points of interest in the subject matter. In other words, diversity of students' academic majors and levels are conditions which govern the range and character of enrollment motives and subject matter interests, respectively.

When the range of enrollment motives for a course is great, it will necessarily encompass reasons which are totally unrelated to instructional effectiveness; thus, satisfaction of interests and student ratings will be lower than if the range were small and restricted to reasons related directly to the nature and quality of instruction. Similarly, all other things being equal, when the range of salient interests in the subject

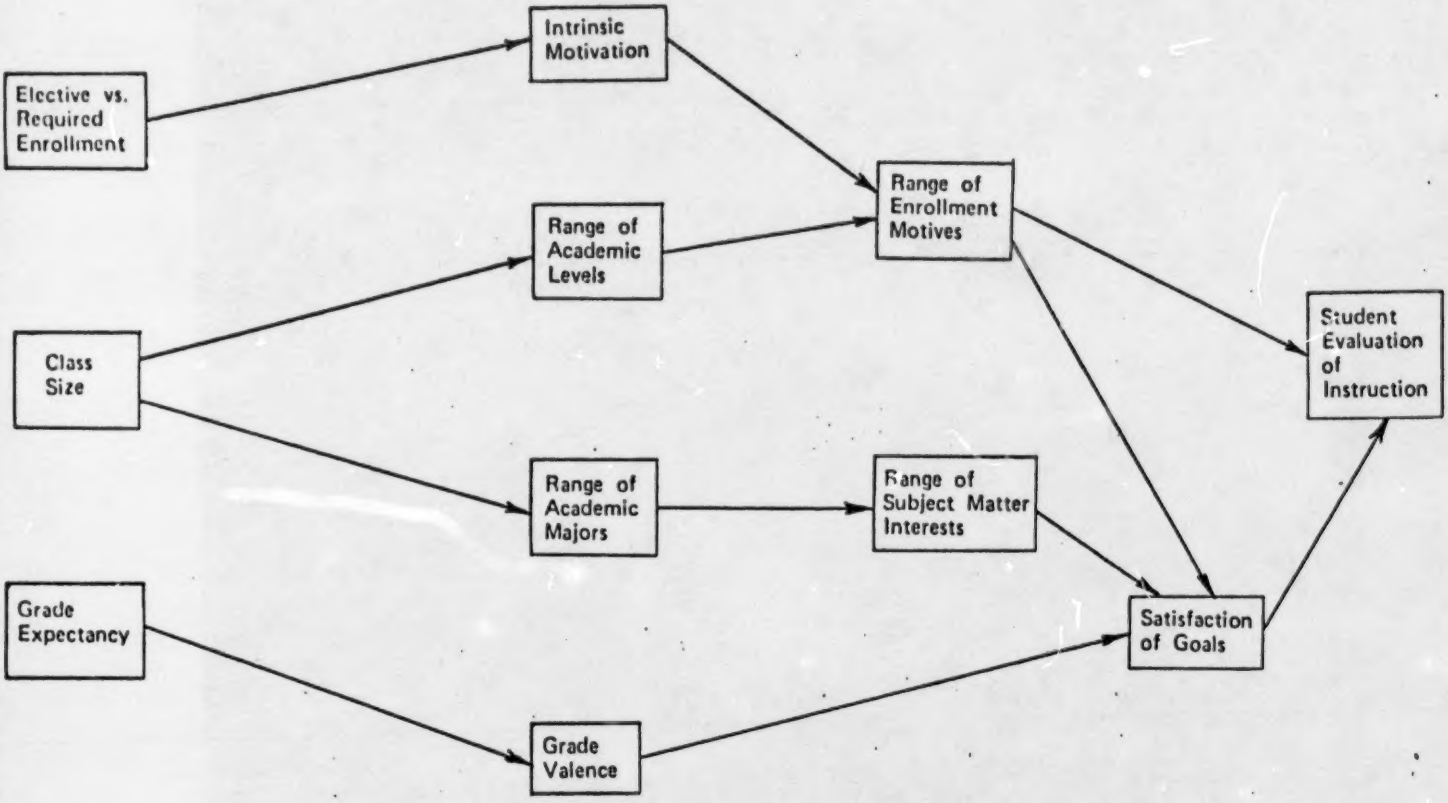

Figure 1. Diagrammatic representation of proposed theory

matter is large (i.e., encompassing professional, personal, academic, and other concerns) it is more difficult to match these interests to the satisfaction of students. Satisfaction of interests or goals will determine, at least in part, student evaluations.

The previous assertions make up the core of the interpretive framework tested. In the interest of a comprehensive examination of variables, however, the roles of elective enrollment and expected grade should also be studied. It was proposed that elective versus required enrollment, as usually quantified, is only a poor indicator of whether enrollment was motivated intrinsically or extrinsically. Moreover, intrinsically motivated students should be more likely to restrict their enrollment motives to those which are under the control of teachers. Finally, expected grade, particularly as it relates to a student's past performance (cumulative grade point average), is highly related to satisfaction of goals.

Method

Subjects

Subjects for this study were students enrolled in 41 undergraduate and 11 dual-level courses at Purdue University during the Spring 1978 semester. In total, data were collected from 1939 students. However, because all variables required measurement of entire classes, individual course sections (N = 52) served as units of analysis. These classes were taught by 43 different instructors.

Because the situational variable Class Size was of considerable importance to the theory tested in this study, a quota sampling procedure

(Kerlinger, 1973) was undertaken in order to derive a sample which might best generalize to the university population. Information obtained from the Purdue University Office of Schedules and Space showed that classes in the University conform to the following class size distribution: 0-9, 11.26%; 10-19, 20.29%; 20-29, 29.53%; 30-39, 12.89%; 40-59, 12.01%; 60-79, 4.86%; and 80+, 9.16%. Sample frequencies in these intervals as well as expected values based on normative percentages are presented in Table 1. A chi-square test for goodness-of-fit (Lynch and Huntsberger, 1976) for this class size distribution yielded a value of 1.138 ($df = 6$). The value is not significant at the $\alpha = .05$ level. From this, it can be concluded that the sample approximates, within reasonable limits of error, the University distribution with respect to class size.

Although not as vigorously controlled as the previously detailed aspect, academic disciplines are also well-represented in the sample. In descending order, discipline representation in the 52 classes is as follows: General Education, 15; Psychology, 8; Chemistry, 7; Spanish, 5; English, 4; Art and Design, 2; History, 2; Health and Safety, 2; Industrial Education, 2; Sociology, 2; Nursing, 1; Electrical Engineering Technology, 1; Anthropology, 1.

Instrumentation

For the purpose of this research, the term Motivational Construct Survey (MCS) will be used to refer to the instrument with which data

Table 1

Class Size Sampling Intervals, Expectancies, and Actual Sample Frequencies

Interval	Percentage[a]	Expectancy	Frequency
0 - 9	11.26	5.86	5
10 - 19	20.29	10.55	9
20 - 29	29.53	15.36	15
30 - 39	12.89	6.70	8
40 - 59	12.01	6.25	7
60 - 79	4.86	2.53	2
80 +	9.16	4.76	6
TOTAL	100.00	52	52

[a]Source of interval percentages was the Purdue University Office of Schedules and Space.

for all variables were collected. Forms responded to by subjects, however, were entitled simply "Information Survey." It consisted of five parts:

I. Personal Information

1. Course required or elective
2. Expected grade
3. Academic level
4. Cumulative grade - point average (GPA)

II. Core Ratings from the Purdue Instructor and Course Appraisal (CAFETERIA; Derry, Siebert, Starry, Van Horn, and Wright, 1974). Five general items rated Strongly Agree (5) to Strongly Disagree (1).

1. My instructor motivates me to do my best work.
2. My instructor explains difficult material clearly.
3. Course assignments were interesting and stimulating.
4. Overall this course is among the best I have taken.
5. Overall, this instructor is among the best teachers I have known.

III. Enrollment. A) Ratings on a 3-point scale of the importance of various possible enrollment motives.

1. Friend(s) recommended this course to me.
2. I had heard that this instructor was a good teacher.
3. I thought it important to learn this material.
4. I thought it would be an easy grade.
5. I thought this course would look good on my transcript.
6. I took this course because my academic program (plan of study) says I have to.

7. I wanted to be with friends.
8. My advisor recommended that I take this course.
9. The organization or method of instruction appealed to me.
10. The time or location of class meetings was more convenient than others.

(B) Rating on a 5-point scale of agreement to the statement, whether this course was required or elective, at the beginning of this semester, I wanted to take it anyway.

IV. Subject Matter Interests. Ratings on a 3-point scale of the degree of interest of various possible areas.

1. A long-time personal interest in the subject matter.
2. Curiosity about course topics.
3. Professional interest in the subject matter.
4. To learn about real-life applications of the material.
5. To learn material useful for later courses.
6. To learn the theoretical foundations of the area.
7. To strengthen a weakness in my background.

V. Involvement and Satisfaction. Subjects were asked to select items from the previous list rated as "very interested" and subsequently rate the degree to which they were addressed, challenged, and/or satisfied. Other items were to be left unrated.

Pilot Study of the MCS

An initial pilot study of the MCS was undertaken to determine the exhaustiveness of lists used in Sections III, IV, and V. Sources for

these lists were items used in the research of Bush (1954), Centra and Creech (1976), and Levinthal, Abrami, Perry, and Breen (1975). Subjects for this study were 33 students enrolled in two sections of an upper-level teaching methods course. A test-retest reliability study (Thorndike, 1951) was conducted. A consumable form of the "Information Survey" was administered during the sixth and eighth weeks of this eight-week course.

The criterion for evaluating the exhaustiveness of enrollment motive and subject matter interest lists was to include any additional items that were supplied by the same subject in both the test and re-test administrations. Additions to the lists were rare; and in no case was an added motive or interest found in both administrations. From this information, then, it was concluded that the lists were indeed exhaustive.

A second pilot study was then undertaken to address the critical issue of whether the MCS elicited responses that were sufficiently reliable to allow a valid investigation of the proposed theory. To meet this purpose, a simple presentation of reliability coefficients and p values will not suffice. The real question was a matter of determining the limits of possible associations with other variables that the precision of MCS data would allow. Therefore, when reliability data are presented, a nominal level of statistical significance ($p < .05$) can be assumed unless otherwise stated. The central question of precision will be addressed.

It must be stated here that although classes were the units of analysis for the main study, individual students were the units for

this pilot study. This was done because insufficient classes were available to allow a class-wise investigation in both main and pilot studies. This is not a serious problem, as class medians are likely to be much more stable than individual responses.

A second draft of the MCS was used in this pilot study. Subjects for Pilot Study 2 were 56 students enrolled in three sections of an introductory educational psychology course. Enrollments in the sections were 25, 14, and 17. Data were collected during the twelfth and fourteenth weeks of the Spring 1978 semester for two of the classes. Stability of responses over a three-week interval was examined for the third group.

The precision of data elicited by the MCS to measure simple and derived situational variables appears to be very high. The choice of elective or required enrollment options did not change over time (r_{xx} = 1.0 for 2- and 3-week intervals). Stability coefficients for expected grade were extremely large (.90 for two-week interval, .88 for three weeks).

Academic levels and departmental affiliations were reported by students in both administrations without any discrepancies. The ranges of these respective student characteristics are, therefore, seen as being very stable. Stability coefficients for the item that measured intrinsic motivation (Section III, B) were .71 (two weeks) and .93 (three weeks). Finally, test-retest correlations for the difference between expected grade and reported GPA were .80 (two weeks) and .90 (three weeks).

Reliability coefficients for individual judgments of the importance of enrollment motives ranged from .46 to .89 (two weeks). Over a three week interval the range was .71 to .93. Clearly, the MCS elicits reasonably stable judgments in this respect. Reliability coefficients for ratings of subject matter interests (Items 15-21) ranged from .74 to .90 (two-week interval) and from .81 to .95 (three weeks).

The range of stability coefficients for satisfaction items (22-28) over a two-week interval was .30 (n = 24, p = .075) to .78. The average reliability of these items (after Z transformation) was .63. The range for the three-week interval group was from .55 (n = 9, p = .06) to 1.0 (n = 9). The average reliability coefficient for these items was .93. Using the more conservative estimate of overall stability, aggregate satisfaction scores have a potential for correlating with other variables no more than r = .79 (square root of the original value). Given the considerable precision of situational variables, the effective ceiling is probably not much lower.

Procedures

Copies of the MCS and mark-sense cards were distributed in each class by the experimenter or a proctor, not the instructor. Subjects were asked to complete the form, following MCS directions. Data were gathered in all classes on the day of regular student evaluations and during the last two weeks of the Spring 1978 semester.

Study Variables

Variables that comprise the proposed theory were computed from responses of subjects to the MCS and from other sources. Operational definitions are as follows:

Class Size: The number of students enrolled in the class (taken from the roster).

Grade Expectancy: Median expected grade for a class (six-point scale).

Elective vs. Required Enrollment: The proportion of students enrolled in a class on an elective basis.

Grade Valence: Median value of the differences between student grade expectancies and GPA.

Range of Academic Majors: The number of academic majors (departmental affiliations) represented in a class.

Range of Academic Levels: The number of academic levels (freshman, sophomore, etc.) represented in a class.

Intrinsic Motivation: Median response to the question, "Whether this course was elective or required, at the beginning of this semester, I wanted to take it anyway." (Strongly Agree = 5, Agree = 4, etc.).

Range of Subject Matter Interests: The number of discrete points of interest identified by three quarters or more of a class.

Range of Enrollment Motives: The number of discrete reasons for enrollment identified by at least one half of a class.

Satisfaction of Goals: Median rating by students of areas of extreme interest in terms of perceived satisfaction.

Student Evaluation of Instruction: Factor score of the five CAFETERIA core items. See Table 2 for loadings on Factor 1 and Cronbach alpha reliability.

Data Analysis

Path analysis (Wright, 1921) was employed to test the adequacy of the proposed theory. It is not a method for proving causality. Rather, it is used to identify the implications of a set of causal assumptions which the researcher is willing to impose upon a system of observed relationships (Kim and Kohout, 1975). As the proposed model is a recursive one (i.e., one-way causality) path analysis procedures detailed by Anderson and Evans (1974) were adopted.

Causal ordering for this analysis can be inferred by the direction of proposed effects shown in Figure 1. Variables on different "branches" of the model (e.g., Intrinsic Motivation and Grade Valence) are not ordered. Path analysis assumes that the highest-order variables (in this case; simple situational variables) are independent. In fact, only Class Size was manipulated; Elective vs. Required Enrollment and Grade Expectancy were not. Any correlation among these variables is probably due, in large measure, to the sampling procedure. Therefore, an additional ordering which reflects this situation is presented in Figure 2. These bivariate relationships are partly causal and partly spurious (Kim and Kohout, 1975).

Table 2

Factor Analysis of CAFETERIA Core Items for 1939 Students[a]

Item	Mean	S.D.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	h^2 (b)
1	3.49	1.07	.538	.495	.262	.241	.000	.688
2	3.89	.92	.652	.345	.247	.040	.023	.607
3	3.31	1.08	.320	.688	.242	.072	-.013	.640
4	3.20	1.21	.348	.585	.580	.074	.050	.808
5	3.69	1.15	.621	.293	.554	.126	-.058	.797
	Eigenvalve		3.183	.200	.117	.035	-.005	
	Pecent of Variance		90.2	5.7	3.3	.7	.1	
	Cumulative Percent of Variance		90.2	95.9	99.2	99.9	100.0	

[a]Varimax rotated factor matrix after rotation with Kaiser normalization (Nie, Hull, Jenkins, Steinbrenner & Bent, 1975)

[b]Communality

Note: Cronback alpha internal consistency reliability for Factor 1 scores = .866.

Zero-order correlations among study variables and path coefficients were obtained by means of the SPSS-REGRESSION program (Nie, et al., 1975). Standardized (β) regression weights were used in this analysis. This was done because beta weights allow easier interpretation (Wright, 1960). Specifically, a given path is interpreted as the standard deviation change in the dependent variable resulting from a one standard deviation change in the independent variable.

Zero-order correlations among the proposed, causally-related variables were "reconstructed" using the standard path theorem (Spady and Greenwood, 1975). A mathematical representation of the theory is

$$\underline{r}_{ij} = \sum_{q} \underline{P}_{iq} \, \underline{r}_{qj}$$

where $\underline{r}$s are correlation coefficients constructed under the assumptions of the model, i and j are indices of the dependent and independent variables, respectively, and q includes every variable which has a path leading directly to the dependent variable. When the model is fully identified (no deleted paths), original correlation coefficients will be completely reconstructed. However, in restricted models, discrepancies generally occur.

The mathematical adequacy of a restricted model will, in large measure, be judged in terms of the frequency of large discrepancies. Within the context of a pair of ordered variables in a restricted model, the discrepancy (spurious or joint effect) is assumed to be due to their causal dependence on a still-higher ordered variable. Spurious effects should be uniformly small. However, strict guidelines for the evaluation of these values do not exist.

Results and Discussion

Means, medians, standard deviations and standard errors for study variables are presented in Table 3. As stated previously, multiple linear regression analyses (Nie, et al., 1975) were conducted in order to obtain path coefficients for the model. These analyses are summarized in Table 4. The resulting path diagram appears in Figure 2.

Potential Causes of Student Evaluation of Instruction

Table 5 contains values for the observed and potentially causal relationships between Student Evaluation of Instruction and all higher-order variables. Satisfaction of interest-related goals has the largest direct effect on student ratings (.4222). The effect of the Range of Enrollment Motives is predominently direct (.1048), but it is also mediated by satisfaction (.0638). These two values yield a sum which very closely approximates the total association (r = .1938). Multiple R for the two predictors was .4560. The indirect effect of Range of Subject Matter Interests (.1213) is nearly sixty percent of its observed relationship.

Despite a moderate positive correlation between Intrinsic Motivation and student ratings (.3068), the restricted model provides for only a

Table 3

Descriptive Statistics for Study Variables*

Variable	$\bar{X}$	Mdn	s	$s_{\bar{X}}$
Class Size	56.250	27.500	96.256	13.348
Elective vs. Required Enrollment	.328	.188	.280	.039
Expected Grade	5.387	5.385	.427	.059
Intrinsic Motivation	3.740	3.755	.629	.087
Range of Academic Levels	3.115	3.250	1.215	.169
Range of Academic Majors	7.827	6.000	7.677	1.065
Grade Valence	.212	.198	.441	.061
Range of Enrollment Motives	3.962	3.682	1.455	.202
Range of Subject Matter Interests	1.173	.463	1.478	.205
Overall Satisfaction	4.241	4.065	.327	.045
Student Evaluation of Instruction	.126	.196	.428	.059

*N = 52

Table 4

Multiple Correlation Coefficients and Standardized Regression Coefficients for 52 Undergraduate and Dual-Level Classes

Dependent Variable	Independent Variable	Multiple Correlation Coefficient	Standardized Regression Coefficient	Standard Error of Beta
SEI	SG	.4560	.4222	.1288
	REM		.1048	.1288
SG	GV	.4558	.2538	.1320
	RSMI		.2874	.1284
	REM		.1512	.1308
REM	RAL	.3117	-.2258	.1348
	IM		-.1977	.1348
RSMI	RAL	.3895	-.3895	.1290
IM	EvsR	.3494	.3494	.1312
RAL	CS	.3494	.3494	.1312
RAM	CS	.8556	.8556	.0725
GV	GE	.7465	.7465	.0932
EvsR	CS	.3498	.3498	.1312
GE	CS	.1484	-.1484	.1385

Note: SEI = Student Evaluation of Instruction
SG = Satisfaction of Goals
REM = Range of Enrollment Motives
RSMI = Range of Subject Matter Interests
IM = Intrinsic Motivation
RAL = Range of Academic Levels
RAM = Range of Academic Majors
GV = Grade Valence
EvsR = Elective vs. Required Enrollment
GE = Grade Expectancy
CS = Class Size

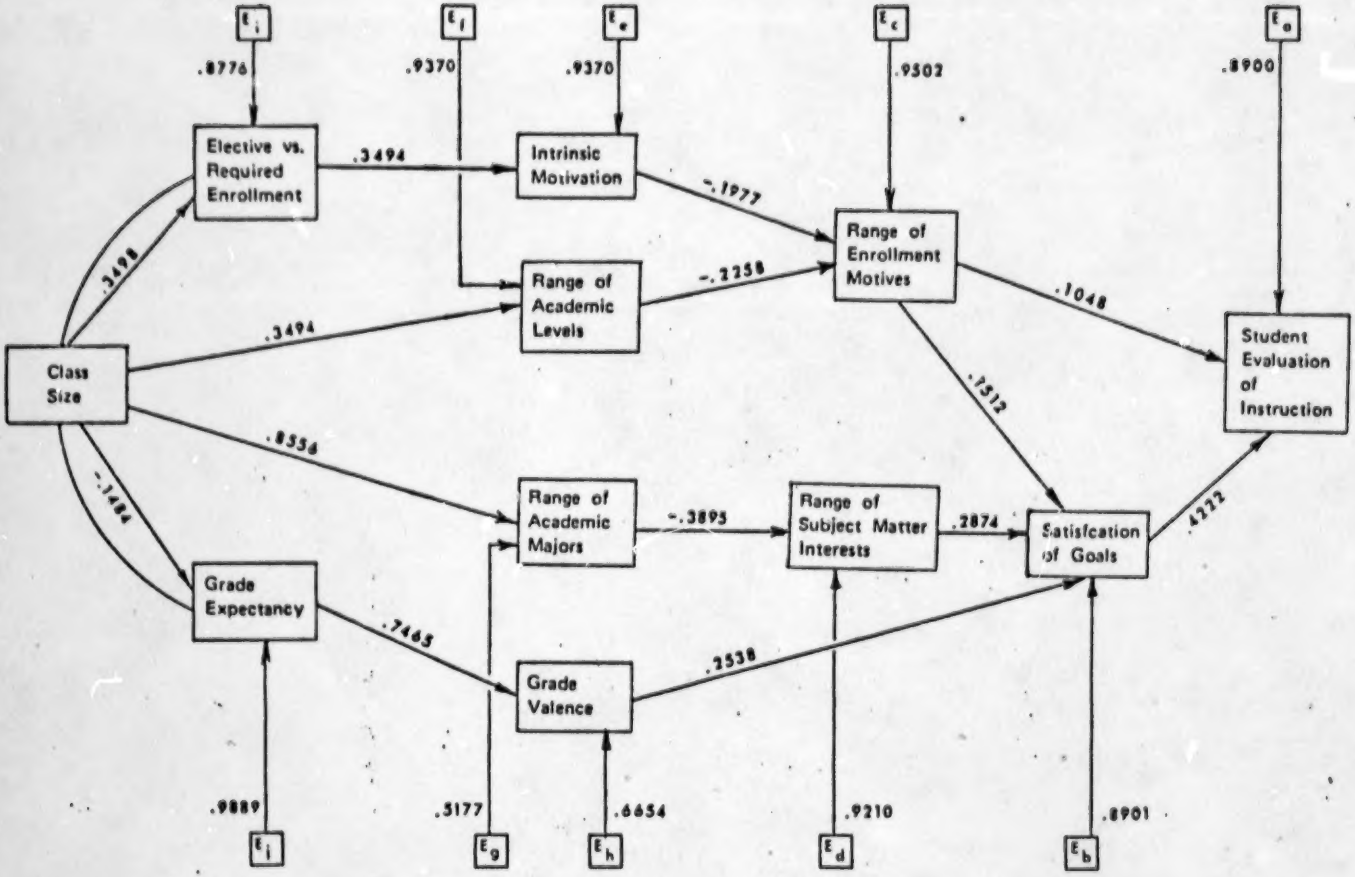

Figure 2. Path diagram for model

Table 5

Contribution of Variables to

Student Evaluation of Instruction

Variable	Total Effect	CAUSAL EFFECTS			Spurious Effect
		Direct	Indirect	Total	
SG	.4443	.4222	.0000	.4222	.0221
REM	.1938	.1048	.0638	.1686	.0252
RSMI	.2058	.0000	.1213	.1213	.0845
IM	.3068	.0000	-.0333	-.0333	.3401
RAL	.0052	.0000	-.0381	-.0381	.0433
RAM	-.1150	.0000	-.0473	-.0473	-.0677
GV	.2544	.0000	.1072	.1072	.1472
EvsR	.1099	.0000	-.0116	-.0116	.1215
GE	.3258	.0000	.0800	.0800	.2458
CS	-.1350	.0000	-.0698	-.0698	-.0652

Note: See Table 4 for variable abbreviations.

small indirect effect (-.0333). The indirect effect of the Range of Academic Levels through the Range of Enrollment Motives is similarly weak (-.0381). The effect of the Range of Academic Majors (-.0473) through the Range of Subject Matter Interests accounts for approximately forty percent of observed covariation. Nearly the same amount was explained for Grade Valence (.1072, through Satisfaction of Goals).

Finally, among the simple situational variable, Elective vs. Required Enrollment was observed to have a weak association with student ratings (.1099). Within the context of the model, however, the potential causal effect is negligible (-.0116). Only twenty-five percent of the observed association between Grade Expectancy and ratings was found to be contained in Grade Valence (.0800). Slightly more than half of the observed correlation between ratings and Class Size was mediated by lower-order variables.

An implication of the restrictions of the model is that some disturbingly large discrepancies between observed and reconstructed correlation coefficients are indeed spurious effects. The most troublesome of these are the values associated with Intrinsic Motivation and Elective vs. Required Enrollment.

The effects of both variables were hypothesized to be contained entirely in Range of Enrollment Motives. However, it was further anticipated that the direction of the path value for this variable would be negative. The model implies that when the range is small, salient motives will involve the reputation of the teacher or the method of instruction. In fact, this was not the case. Fifty-one sample classes identified the

reason, "I thought it important that I learn this material," as being at least of minor importance. "My advisor recommended that I take this course" was selected by 44 classes. "I took this course because my academic program (plan of study) says I have to" was an important motive for 41 classes. The sample median for this variable is 3.68. Clearly, when the range of enrollment motives is small, they are restricted, for the most part, to these three reasons. A positive path coefficient indicates that high student ratings tend to be associated with classes for which reasons for enrollment extend beyond these common three.

Potential Causes of Satisfaction of Goals

Table 6 contains values for the observed and reconstructed correlations between Satisfaction of Goals and all higher-order study variables. Ranges of Enrollment Motives and Subject Matter Interests had direct paths which were very close to their observed correlations (.1512 and .2874, respectively). The direct path for Grade Valence (.2538) was more than three quarters of the observed relationship. Multiple R between these three predictors and satisfaction of interest-related goals was .4558.

As in the previous section, despite a moderate positive correlation between Intrinsic Motivation and Satisfaction Goals (.3206), the restricted model provides for only a small indirect effect (-.0299). The indirect effect of the Range of Academic Levels contained in the Range of Enrollment Motives is also small (-.0341). The effect of diversity in academic majors (-.1119), mediated by the Range of Subject

Matter Interests, accounts for nearly all of the original correlation.

Again, among the situational variables, the proportion of elective enrollments in the classes had a weak positive correlation with ratings of satisfaction (.0738). However, the model shows only a negligible indirect effect through Intrinsic Motivation (-.0104). Only one-third of the association between Grade Expectancy and satisfaction was found to be contained in the direct effect of Grade Valence (.1895). Finally the effect of class size, mediated through lower-order variables (-.1395) is very close to the observed correlation.

Again, the implication of this aspect of the analysis is that observed correlations of Intrinsic Motivation and Elective vs. Required Enrollment with Satisfaction of Goals are totally spurious.

The reason for such large discrepancies is again the direction of correlation between the dependent variable and Range of Enrollment Motives. The positive path value indicates that high levels of satisfaction tended to be present in classes for which salient motives extended beyond simple program requirements.

Potential Causes of the Range of Enrollment Motives

Table 7 contains values for the observed and reconstructed correlations between Range of Enrollment Motives and all lower-order variables. Direct paths for Intrinsic Motivation (-.1977) and Range of Academic Levels (-.2258) are very close to their respective zero-order

Table 6

Contribution of Variables to

Satisfaction of Goals

Variable	Total Effect	CAUSAL EFFECTS			Spurious Effect
		Direct	Indirect	Total	
REM	.2109	.1512	.0000	.1512	.0597
RSMI	.3223	.2874	.0000	.2874	.0349
IM	.3206	.0000	-.0299	-.0299	.3505
RAL	.0569	.0000	-.0341	-.0341	.0910
RAM	-.1033	.0000	-.1119	-.1119	.0086
SV	.3279	.2538	.0000	.2538	.0741
EvsR	.0738	.0000	-.0104	-.0104	.0842
GE	.5744	.0000	.1895	.1895	.3849
CS	-.1245	.0000	-.1395	-.1395	.0150

Table 7

Contribution of Variables to

Range of Enrollment Motives

Variable	Total Effect	CAUSAL EFFECTS			Spurious Effect
		Direct	Indirect	Total	
IM	-.2155	-.1977	.0000	-.1977	-.0178
RAL	-.2415	-.2258	.0000	-.2258	-.0157
EvsR	-.1958	.0000	-.0691	-.0691	-.1267
CS	.0611	.0000	-.1031	-.1031	.1642

Note: See Table 4 for variable abbreviations.

correlations. Multiple R for these predictors was .3117.

The indirect effect of Elective vs. Required Enrollment (-.0691), through Intrinsic Motivation, was slightly more than one-third the observed association. Class size had a very small correlation with Range of Enrollment Motives (.0611). However, within the context of the model, its indirect effect through the Range of Academic Levels is negative and somewhat stronger (-.1031).

Because Class Size is the highest-order variable, a large spurious effect threatens the assumption of causal closure. However, the observed correlation is very small; and the standard error for any value computed on the basis of 52 cases is relatively sizable (.14). This discrepancy, therefore, is not seen as a serious threat to the assumption of causal closure.

Potential Causes of the Range of Subject Matter Interests

Table 8 contains values for the observed and reconstructed correlations of Range of Subject Matter Interests with Range of Academic Majors and Class Size. Because the Range of Academic Majors was the only proposed determinant of the variety of interests, no discrepancy was found between the direct path and original correlation values (-.3895). The indirect effect of Class Size contained in academic majors was -.3333. This negative effect is stronger than the observed relationship (-.2551) by some thirty percent.

Table 8

Contribution Variables to

Range of Subject Matter Interests

Variable	Total Effect	CAUSAL EFFECTS			Spurious Effect
		Direct	Indirect	Total	
RAN	-.3895	-.3895	.0000	-.3985	.0000
CS	-.2551	.0000	-.3333	-.3333	.0782

Note: See Table 4 for variable abbreviations.

Potential Causes of Intrinsic Motivation

Table 9 contains values for observed and reconstructed correlations of Intrinsic Motivation with Elective vs. Required Enrollment and Class Size. Because the proportion of elective enrollees was the only proposed determinant of Intrinsic Motivation, no discrepancy between the two values was found ($\underline{P}$ = .3494). Class Size was observed to have little relationship with Intrinsic Motivation ($\underline{r}$ = .0128). However, within the context of the model, its indirect effect through Elective vs Required Enrollment is .1222.

Neither value exceeds the standard error for a sample of 52 cases. Moreover, the ordering of Class Size and Elective vs. Required Enrollment was due to the lack of control over the variation of the latter and not to theoretical considerations. Therefore, this discrepancy is not seen as a threat to the assumption of causal closure.

Potential Causes of Grade Valence

Table 10 contains values for observed and reconstructed correlations of Grade Valence with Grade Expectancy and Class Size. Because expected grade was the only proposed determinant of Grade Valence, no discrepancy between the two values was found (P = .7465). Class size was observed to have an extremely small relationship with Grade Valence ($\underline{r}$ = .0451). But within the context of the model, its indirect effect through Grade Expectancy is -.1108. Again, neither

Table 9

Contribution of Variables to

Intrinsic Motivation

Variable	Total Effect	CAUSAL EFFECTS			Spurious Effect
		Direct	Indirect	Total	
EvsR	.3494	.3494	.0000	.3494	.0000
CS	.0128	.0000	.1222	.1222	-.1094

Table 10

Contribution of Variables to

Grade Valence

Variable	Total Effect	CAUSAL EFFECTS			Spurious Effect
		Direct	Indirect	Total	
GE	.7465	.7465	.0000	.7465	.0000
CS	.0451	.0000	-.1108	-.1108	.1559

Note: See Table 4 for variable abbreviations.

value exceeds its own standard error. And because the ordering of Class Size and Grade Expectancy was due to the lack of control of expectancies in the sampling procedure, the threat to causal closure is not serious.

Class Size as a Potential Cause of Other Situational Variables

Class Size was the only proposed causal agent for the variables. Range of Academic Levels, Range of Academic Majors, Elective vs. Required Enrollment, and Grade Expectancy. Therefore, path values were equal to the original zero-order correlations (see Figure 2) and no discrepancies are reported.

Summary of Path Analysis

A mathematical summary of the theory is presented below in the form of structural equations. Variable abbreviations are used and indicate that they are in standard form.

$$SEI = .4222(SAT) + .1048(REM) + .8900(E_a)$$

$$SAT = .1512(REM) + .2874(RSMI) + .2538(GV) + .8901(E_b)$$

$$REM = -.1977(IM) - .2258(RAL) + .9502(E_c)$$

$$RSMI = -.3895(RAM) + .9210(E_d)$$

$$IM = .3494(EvsR) + .9370(E_e)$$

$$RAL = .3494(CS) + .9370(E_f)$$

$$RAM = .8556(CS) + .5177(E_g)$$

$$EvsR = .3498(CS) + .8776(E_i)$$

$$GE = -.1484(CS) + .9889(E_j)$$

Summary and Conclusions

The purpose of this study was to test a theory for the potential effects of situational variables on student evaluation of instruction. It was hypothesized that simple situational variables (class size, elective enrollment, expected grade) would be less potent determinants of student ratings than certain derived situational variables (diversity of academic levels and majors, intrinsic motivation and grade valence). Grade valence is the difference between the expected grade and cummulative grade point average. These variables, in turn, were hypothesized to give way to even more effective predictors, proposed motivational constructs (diversity of enrollment motives and subject matter interests, and the rated satisfaction of important interests). The basic assertion of this theory is that situational variables are not immediately related to student ratings. However, they do set the conditions under which these motivational constructs take on values.

Path analysis of the total associations (zero-order correlations) among study variables generally supported this notion of a hierarchy of determinants for student ratings. However, specific hypotheses, which operationalized this theory, were not all supported. In particular, it was hypothesized that when the rated level of intrinsic motivation is high and when the range of academic levels is narrow, the salient reasons for enrollment will be restricted to the reputation of the instructor and the organization of the course. The result of the salience of these two motives was then hypothesized to be a tendency towards higher ratings (i.e., an inverse relationship between range of motives and student ratings).

In fact, it was found that in almost all sample classes, program requirements and advisor recommendation were important motives. Thus the aforementioned reasons were salient only when the range was relatively large. Observed relationships between Range of Enrollment Motives and both Satisfaction of Goals and Student Evaluation were in the opposite direction anticipated. The result of this unexpected finding was an apparent lack of representation of the positive relationships between intrinsic motivation and elective enrollment and lower-order variables in the model.

The failure to reproduce correlations of ratings with intrinsic motivation and elective enrollment suggests that the accepted model may not be causally closed as had originally been assumed. Similarly, less than fifty percent of the observed associations with grade expectancy and valence are explained by the model. This is despite the fact that their proposed effects are contained in the strongest predictor of rating variation (i.e., Satisfaction of Goals). The threat to causal closure, however, is uncertain because the variability of Grade Expectancy and Elective vs. Required Enrollment were not controlled in the sampling procedure as was Class Size. Therefore, unusually large correlations may have resulted. In addition, end-of-course ratings of initial intrinsic motivation might have also been influenced by the evaluation setting. These caveats notwithstanding, the model as originally proposed offers some important implications for the use of ratings in various contexts. The remainder of this article is devoted to these implications.

Despite the recent proliferation of rating systems in post-secondary institutions, widespread acceptance of evaluation data to suit a variety of purposes has been lacking. A major reason for past skepticism has been that student ratings are supposedly subject to a host of situational influences. A common notion is that a sure-fire remedy for low ratings is to teach small, elective classes and to give all "A"s. The findings of this study dispute that contention.

Path analysis has shown that, under the restrictions imposed by the theory, the relationships between ratings and the size of classes as well as diversity of academic levels and majors are completely contained in the direct effects of certain psychologically-oriented variables. This is to say that any potentially inverse effect of class size or resulting diversities of levels and majors is not manifested immediately upon student ratings. The more plausible explanation is that large classes of diverse students tend to see their enrollment as satisfying externally motivated pressures (i.e., program requirements, advisor referral). Similarly, a tendency towards a lack of common interests was observed. Therefore, sources of personal satisfaction are limited. Each of these outcomes was shown to have a potentially negative effect on student ratings.

Implications for the diagnostic use of ratings can be derived directly from observed determinants. Although not overpowering, the most important determinant of student evaluation is the satisfaction of important subject matter interests. Clearly, teachers who wish to improve their instructional effectiveness (Ratings should be an afterthought.) should identify salient student interests and design instruction

to meet them. Reasons for enrolling in a course and information about past achievement can also be useful when planning instruction.

The common thread is individual differences. Just as students differ in terms of their general preparedness for a course, they also differ in terms of motives and interests. This study has shown that student ratings are more sensitive to these differences than to situational and student demographic differences across classes.

References

Anderson, J.G., & Evans, F.B. Causal models in educational research: Recursive models. American Educational Research Journal, 1974, 11, 29-39.

Bush, R. The teacher-pupil relationship. New York: Prentice-Hall, 1954.

Centra, J.A. & Creech, F.R. The relationship between student, teacher, and course characteristics and student ratings of instructional effectiveness. (ETS PR-76-1). Princeton: Educational Testing Service, 1976.

Clark, K.E. & Keller, R.J. Student ratings of college teaching. In R.A. Eckert (Ed.), A university looks at its program. Minneapolis: University of Minnesota Press, 1954.

Cohen, J. & Humphreys, L.G. Memorandum to faculty. University of Illinois, Department of Psychology, 1960. (Mimeographed)

Costin, F., Greenough, W.T., & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Derry, J. O., Siebert, W.F. Starry, A.R., Van Horn, J. W. & Wright, G. L. The CAFETERIA System: A new approach to course and instructor evaluation. West Lafayette, Indiana: Purdue Research Foundation, 1954.

Downie, N.M. Student evaluation of faculty. Journal of Higher Education, 1952, 23, 495-496, 503.

Feldhusen, J. F. & Mazzuca, S.A. Effects of situational and performance variables on instructional ratings. College Student Journal, in press.

Frey, P.W. A two-dimensional analysis of student ratings of instruction. (Unpublished research) Northwestern University, 1977.

Heilman, J. D. & Armentrout, W.D. The rating of college teachers on ten traits by their students. Journal of Educational Psychology, 1936, 27, 197-216.

Kerlinger, F. N. Foundations of behavioral research. New York: Holt, Rinehart & Winston, Inc., 1973.

Kim, J. & Kohout, F. J. Special topics in general linear models. In N. H. Nie, C.H. Hull, J. G. Jenkins, K. Steinbrenner, & D.H. Bent (Eds.) SPSS Statistical package for the social sciences. New York: McGraw-Hill, 1975.

Kulik, J. A. & McKeachie, W.J. The evaluation of teachers in higher education. In F.L. Kerlinger (Ed.), Review of Research in Education 1975, 3, 210-240.

Levinthal, L., Abrami, P., Perry, R. & Breen, L. Section selection in multisection courses: Implications for the validation and use of teacher ratings forms. Educational and Psychological Measurement, 1975, 35, 885-895.

Lovell, G. D. & Haner, C. F. Force-choice applied to college faculty ratings. Educational and Psychological Measurement, 1955, 15, 291-304.

Lynch, M. D. & Huntsberger, D. V. Elements of statistical inference for education and psychology. Boston: Allyn & Bacon, 1976.

McDaniel, E. D. & Feldhusen, J. F. Relationships between faculty ratings and indexes of service and scholarship. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 619-620.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. & Bent, D.H. SPSS-Statistical package for the social sciences, (2nd. ed.). New York: McGraw-Hill, 1975.

Rayder, N. F. College student ratings of instructors. Journal of Experimental Education, 1968, 37, 76-81.

Spady, W. G. & Greenwood, D. Instant path analysis: An elementary cookbook. Division Generator, 1970, 1, 3-8.

Thorndike, R. L. Reliability, In E. F. Lindquist (Ed.), Educational measurement. Washington: American Council on Education, 1951, 560-620.

Treffinger, D. J. & Feldhusen, J. F. Predicting students' ratings of instruction. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 621-622.

Wright, S. Correlation and causation. Journal of Agricultural Research, 1921, 20, 557-585.

Wright, S. Path coefficients and path regressions: Alternative or complementary concepts? Biometrics, 1960, 16, 189-202.